AUDITORY TESTS.[1]

BENJAMIN RICHARD ANDREWS, A. B., A. M.

TABLE OF CONTENTS.

INTRODUCTION.

The two great functions of the ear are these: it furnishes the most common means of intercommunication, that of spoken speech; and it opens to mankind one whole department of the fine arts, that of music.

Efficient hearing is indispensable if the individual is to do his full part in the world's work, or to enjoy his share of its æsthetic pleasures. Normal functioning of the auditory organs is, therefore, a matter of importance to the individual. This paper discusses tests which may be applied to an individual to determine his auditory capacity as normal or abnormal. Such tests form part of what are commonly called "mental tests," *i. e.*, tests which have to do with mental factors. Our subject has, therefore, a certain relation to psychology. Moreover, since certain auditory tests are used to diagnose aural disease, it has also a relation to pathology.

The first chapter contains an examination of the concept of "mental tests," and of the relation of such tests to psychology and pathology. In the second chapter, the literature of mental tests is briefly reviewed. These two chapters serve as a pref-

---

[1] From the Psychological Laboratory of Cornell University.

ace to the description of the various auditory tests, which forms the main problem of the paper. These tests are treated in three divisions, forming chapters three, four and five: Chapter III: Tests of auditory acuity (a) with spoken words and (b) with mechanical sounds; Chapter IV: Tests of musical capacity; Chapter V: Tests used in diagnosis of aural disease. The present paper includes Chapters I-III; Chapters IV and V will follow in a subsequent article.[1]

CHAPTER I.

Mental Tests and their Relation to Psychology and Pathology.

The term "test" implies the existence, or possibility, of an average quantitative expression for a physical or mental characteristic or capacity, to which as a standard or norm the measurements of given individuals may be referred. The standard or norm value is taken as the average or mean of the measurements of many individuals with regard to the characteristic concerned. This concept lies at the basis of anthropometry.

This science has confined itself thus far almost entirely to physical measurements of the human body. Anthropometrists have compiled tables which show the normal height, weight, strength, lung capacity, length of arm, and other purely physical characteristics, of the typical individual, in various classes arranged according to age, sex, race, occupation and other differentiae. The normal proportions between various measurements, as those of girth and height, have also been recorded. Such norms are ideal values secured by taking the average or mean value of the particular measurements in a large number of individuals of a definite class; they represent the development which the normal individual within the class is most likely to possess. When norms are once established, the measurements of given individuals can be referred to them, and the individuals classed as normal, or as sub-normal or super-normal in a certain degree, with respect to the particular characteristics. The results of physical anthropometry are valuable, first of all, as scientific facts: like the facts of astronomy, they have a value in and of themselves. They find a considerable practical application as well. Successive testing of the same individual measures the influence of exercise, employment, or habits

[1]The writer wishes to express his obligation to Professor I. Madison Bentley, of the Department of Psychology, Cornell University, for putting at his disposal the contents of an unpublished paper on tests of audition, which proved very valuable in giving a starting point for the writer's own investigations.

of life, on his physical organism. Anthropometrical examinations guard the entrance to certain occupations, as the army and police, and in some instances admission into public schools. Such measurements have been adapted by A. Bertillon to furnish an accurate system of identification of criminals. They have recently been introduced into public schools to detect abnormal bodily development and make physical training effective; and their results are throwing light on other pedagogical problems. In all such measurements, there is assumed an average typical development for members of each class, and individual variations from this norm are regarded as of significance. In these physical measurements, which comprise thus far the great bulk of anthropometrical investigations, there is coming to be general agreement as to the dimensions and proportions which shall be noted, their normal values, and, to some extent, the significance of deviations from these values.

Anthropometry has shown also a tendency to use tests involving mental factors and, indeed, to measure mental characteristics themselves. Francis Galton included certain mental measurements in the anthropometrical investigations which he suggested in 1882[1] and afterward carried out. In 1890 the matter was brought to the attention of psychologists by Cattell's article in *Mind*[2] on "Mental Tests and Measurements." He had already applied measurements to typical mental characteristics, and in this article advocated the adoption and wide use by psychologists of tests of vision, audition, reaction time, and other mental phenomena. Other investigators took up the matter, and a few years later, in 1895, the American Psychological Association appointed a committee to select a series of tests through which there could be established "the normal capacity of simple and typical sensory, motor and intellectual endowment in the average individual and in groups; the distribution of these capacities, their correlation, development, and relation to daily activities."[3] The tests were to be used in psychological laboratories upon large numbers of college students and other individuals, and in this way data secured "for the study of general, typical, characteristic endowment, much as the anthropometrist desires first to establish standards of the principal physical measurements and proportions."[4] Such tests were instituted at Columbia, Wisconsin, and some other universities. A programme closely related to that of "mental tests" is the individual psychology proposed in France

[1] *Fortnightly Review*, xxxvii, 1882, 332.
[2] *Mind*, XV, 1890, 373.
[3] J. Jastrow (member of committee): *Psychological Review*, V, 1898, 172.
[4] *Ibid.*

by Binet and Henri and in Germany by Kraepelin and Oehrn. Its problem, as stated by Binet and Henri is (1) the study of the variable properties of psychical processes,—how and to what extent processes vary from individual to individual; and (2) the discovery of the interrelations of these variable properties in the individual mind. Without reviewing the literature of either mental tests or individual psychology at this point, we may say that each proposes to measure such mental functions as memory, association, reaction-time, imagination, audition, vision, attention, etc., in large numbers of individuals; to determine norms as physical anthropometry has done; and to discover and interpret individual variations from them. Each, moreover, expects to contribute to the science of psychology. It will, therefore, be desirable to determine at once how mental tests are related to psychology and whether they come within its scope as a science.

Now modern psychology may take three more or less clearly defined attitudes toward its subject matter, mental experiences. (1) It may regard mental stuff statically, as structure, disregarding its reference in the environment and its significance to the individual. It then finds its immediate problem in the analysis of consciousness into structural elements, the sensations and affections, and in the statement of the laws governing their interconnection in the structural complexes found in consciousness. (2) It may regard the processes dynamically, as functions of the psychophysical organism, having meaning to the organism in its relations to the outside world. The problem, as before, is, first, analysis, but now into elementary functions; and secondly, a statement of the laws governing their interaction. (3) It may regard mind as in process of development and trace its progress genetically in the race or in the individual. Since this genetic attitude may be either structural or functional, the primary psychological standpoints reduce to two. To which of these are mental tests related? In general, not to structural psychology, we may say, since one tests the individual in some mental characteristic regarded in its significance to him and without regard to its structural aspects, the sensations and affections which make it up. The answer is not so easy regarding the relation of mental tests and functional psychology. They work upon common material, namely, mental functions. They may be concerned with the same function. Yet, from the standpoint of scientific methodology, with which the decision must finally rest, mental tests and functional psychology lie outside each other, and mental tests, as tests, are not a part of the science of pure psychology. Psychology approaches its problem with trained observers, delicate apparatus, careful procedure designed to check errors, rigid experimental conditions,

and with the single purpose of seeing and describing the phenomenon as it is. A mental test of the same function involves the examination of untrained subjects, under limitations of time which make refined measurement and precise procedure impossible, and with the subjective conditions of tranquility, attention, and knowledge of the problem varying widely from observer to observer. Norms are based upon such measurements, and the individual's variation from them is interpreted for some practical end, perhaps medical diagnosis, or a characterization of the individual as of normal capacity in a certain direction. Psychology, on the other hand, does not recognize individuals as such. Just as it is not a fact in the science of theoretical physics that a certain engine produces a certain amount of power per fuel unit, so it is not a problem for psychology to evaluate John Smith's memory or attention. The former problem falls to the province of engineering, an application of pure physics; the latter, a typical mental test, is just as certainly an application of psychology, and is not to be regarded as within its limits as a theoretical science. Further, psychology of function, as a science, has no room for the practical, extra-scientific ends involved in mental tests; it measures mental processes simply to make possible a complete statement of the functional elements of mind and their interaction.

This distinction, which puts mental tests outside psychology, does not exclude the possibility of a scientific individual psychology, which, as Henri and Binet propose, shall describe the variations of mental processes from mind to mind. The objection is simply against the attempt to found such a psychology on the rapid, necessarily inaccurate, and wholesale testing of miscellaneous individuals. The painstaking determinations of trained psychologists, working with refined methods and selected observers, may alone be expected to build up a trustworthy body of facts for a psychology of the individual. The progress of child psychology is illuminating and admonitory in this connection. The established facts regarding the child's mind have been secured by a half-dozen trained investigators through long-continued observations of children; while the mass of facts secured by the questionary and the untrained observations of parents and teachers are largely either unreliable or outside of psychological science. So, doubtless, the development of functional psychology, whether individual or general, will depend on the investigations of experienced psychologists, working with practised observers, and not on the statistical treatment of measurements made on large numbers of individuals taken at random. Mental tests, as tests, therefore, are not to be regarded as part of psychology, neither should direct contributions to psychology be expected from them. Psychology, on the other

hand, can do a great deal for mental tests. In the first place, it must guide in selecting the tests to be employed; and, secondly, it must furnish appliances and methods of procedure, to be adapted as the grosser work of the tests may require.

In concluding this discussion, then, we may say that mental tests, as tests, are more akin to physical anthropometry than to psychology. They are to be sharply distinguished from the measurement of mental functions in psychology, both because of the differences in refinement of method, and particularly because they are applied to masses of individuals to determine norms, individual variations from which are interpreted with regard to some practical end. Mental tests may be considered as an application of psychology, and in so far they are related to it. In the methodology of science, it seems proper to term them, as has been suggested, mental anthropometry.

There remains the further question of the relation of mental tests to pathology, in particular to medical diagnosis. In discussing the applications of physical anthropometry, we mentioned that it has been found useful, as in gymnasiums, in detecting poor physical development, in indicating the exercise or other treatment needed, and in measuring the response of the physical organism to such treatment. Certain tests involving mental factors are used in a very similar way to determine the presence of diseased or abnormal conditions, as the tests of vision, audition, and the rough tests of intelligence to determine insanity. The oculist uses a series of tests whose results indicate primarily the normal or abnormal condition of the visual organs, and, further, if the eye be abnormal or diseased, reveal the nature of the disease or abnormality. Similiarly, the aurist employs tests which indicate whether hearing is sound and normal, and certain other tests which, in case of diseased or abnormal conditions of the ear, reveal the origin, nature, and extent of the disturbance. In such tests, at least so far as diagnosis of disease is involved, we are in the field of pathology. Because of their significance in identifying diseased conditions, they may be called diagnostic tests. While such determinations are primarily matters of pathological practice, yet they involve mental factors, and so become of interest not only for their own sake, but for their relation to psychology. In any case, they should be included, to make our systematic presentation complete. In our treatment of auditory tests, we shall accordingly discuss, first, tests measuring the normal condition of the auditory organ and variations from the normal; and, secondly, certain tests useful to otology, the special branch of pathological science treating of aural disease.

CHAPTER II.

LITERATURE.

This sketch of the literature of mental and auditory tests will not attempt an exhaustive review of all that has been written on the subject. The general bearings of both mental tests and individual psychology, and their literature down to 1898, have been reviewed by Miss Sharp in her article on Individual Psychology;[1] and the literature of auditory tests up to ten years ago, has been summarized by Chrisman.[2] Further, as we come presently to consider various tests of audition, we shall be in constant reference to the literature. For these reasons, a detailed examination of all the literature would be superfluous. At the same time, a brief review of the principal statements of mental tests will be valuable in giving perspective and in relating the present work to that of previous writers, particularly American.

Mental tests, as already stated, go back to Francis Galton. His article on "The Anthropometric Laboratory" in 1882,[3] urged the establishment of laboratories where measurements of purely physical characteristics and of certain mental characteristics as well, could be made. The tests which he suggested include, in addition to purely physical measurements, the following: strength by dynamometer, muscular agility, co-ordination of eye and muscle, sight, hearing, touch, muscular sense, reaction-time, after-images, and memory. Galton's tests of hearing embraced "keenness of hearing, and the appreciation of different degrees of loudness and of different tones." It is to be noted that Galton's standpoint, in suggesting such tests, mental and physical, is that of the anthropometrist. He called attention to the obligation of mental tests to psychology, a debt no one disputes, in his remark that from the work of Fechner and Delboeuf and similar investigators suitable instruments could be selected for "an anthropometric laboratory." He said nothing of possible contributions from mental tests to psychology; this idea was developed, as we shall see presently, by later writers whose interests were in psychology, not in anthropometry. It is important to note that mental tests were first suggested by Galton as anthropometric measurements, quite on a par with physical measurements, without a thought that they would be anything more. This may be fairly inferred from the reasons which he gives for performing such

---

[1] S. E. Sharp: *American Journal of Psychology*, 1899, 329-348. See also her conclusions, pp. 388-391.

[2] Oscar Chrisman: *Pedagogical Seminary*, ii., 1892-3, 397.

[3] *Fortnightly Review*, xxxvii, 332; reprinted in *Popular Science Monthly*, xxi, 53.

tests.[1] It would be interesting to follow Galton's development of the idea of mental tests further; but it must suffice to call attention to his book of 1883, "Inquiries into Human Faculty,"[2] and to the establishment of a temporary anthropometric laboratory in 1884 in connection with the International Health Exhibition at South Kensington, London. At this laboratory, various measurements, some purely physical, some involving mental factors (*e. g.*, the highest audible tone), were made upon over 9,000 persons.[3] The anthropometric laboratory was maintained at South Kensington until about 1895, and then removed to Oxford.

With Galton, mental tests were simply a branch of anthropometry. Their attention at the hands of psychologists seems first to have been due to Cattell. In 1885, and subsequently, he published the first contributions to individual psychology,[4] and in 1890, his article in *Mind*[5] brought the subject of mental tests to the attention of psychologists. Cattell's article was at once an outline of tests already performed, and a suggested programme of tests for general use in the psychological laboratories then forming in American universities. Cattell's tests all involved mental factors, some, as dynamometer pressure, including as well prominent physical factors; others, as memory, being purely mental. It is important to note that his purpose in performing the tests was primarily psychological, to contribute to the science of psychology. He says: "psychology cannot attain the certainty or exactness of the physical sciences, unless it rests on a foundation of experiment and measurement. A step in this direction could be made by applying a series of mental tests and measurements to a large number of individuals. The results would be of considerable scientific value in discovering the constancy of mental processes, their interdependence, and their variation under definite circumstances." While recognizing the practical possibilities of tests, Cattell here apparently puts considerable faith as well in their "scientific value" for psychology. This was in 1890. We may fol-

[1] Galton's reasons are: (1) biographical interest; (2) medical interest to the person tested; (3) information for descendants as to hereditary dangers and vital probabilities; (4) material for investigations into life histories.

[2] Now, unfortunately, out of print.

[3] *Anthropological Institute Journal*, xiv, 1884-5, 205, 275; *ibid.*, xvi, 2. For other references to Galton's work, *cf. Science*, v, 294, and viii, 374; *Nature*, xxxi, 223. *Cf.* also *Lippincott's Mag.*, xlv, 1890, 236, *Why We Measure Mankind*, by Galton.

[4] *Cf.* foot-note, *Psych. Rev.* iii, 1896, 618. Cattell here says that the introduction of individual psychology was delayed because Wundt "was not favorable to it."

[5] J. McKeen Cattell: *Mind*, xv, 1890, 373, *Mental Tests and Measurements*.

low this trend toward a psychological evaluation of mental tests further.

The idea of contributing to psychology by wide-spread testing and measurement seemed to secure general favor. Jastrow, in 1891, published the results of some tests on the community and association of ideas in men and women.[1] In 1892, there were published the results of a series of anthropometrical and psychological tests carried out under Jastrow's direction on students at the University of Wisconsin.[2] The next year, Jastrow conducted a psychological laboratory at the Columbian Exposition in Chicago, in part similar to Galton's anthropometric laboratory, but giving more attention to strictly psychological tests.[3] In the same year (1893), there were two statements of mental tests, by Scripture[4] and Cattell[5] respectively, especially in their relation to school children. The practical value of tests was urged by both, and both seemed impressed with "the scientific value of such tests on large numbers of persons," as Scripture broadly puts it. Cattell, indeed, states explicitly the position with which the writer has taken issue. In urging teachers to perform tests upon school children, he says, "those who wish to contribute to the advancement of psychology will find this a convenient opening." This is a rather radical statement of the relation of tests to psychology, but it seems to represent the prevalent feeling of the time that in mental tests and statistical methods psychology had a new organon. At this juncture, when the proposed psychological laboratory at the Columbian Exposition was under consideration, Titchener published an article on the relations of psychology and mental tests in which there are stated, explicitly

[1] Joseph Jastrow: *A Study in Mental Statistics*, *New Review*, v, 1891, 559; *A Statistical Study of Memory and Association*, *Educational Review*, ii, 1895, 442. Another study by Jastrow was published in 1894: *Community and Association of Ideas: A Statistical Study*, *Psych. Rev.*, i, 1894, 152. This last test was repeated at Wellesley College with different results, under the direction of M. W. Calkins: *Psych. Rev.*, ii, 1895, 363. A criticism of the latter by Jastrow followed: *ibid.*, iii, 1896, 68; to which Miss Calkins replied, *ibid.*, 426. Jastrow made another reply: *ibid.*, 430. The Wisconsin and Wellesley tests are criticised by Amy Tanner: *ibid.*, 548.

[2] *Am. Jour. Psych.*, iv, 1892, 420. The tests were performed in the fall of 1890.

[3] "Experimental Psychology at the World's Fair," an outline of plans given in an address by Jastrow before the American Psychological Association, Philadelphia, December 27, 1892. *Cf. Official Catalogue of the World's Columbian Exposition*, pt. 12, 50. For an analysis of the tests, *cf. L'Année psychologique*, i, 522. No results, as far as the writer has been able to learn, have been published.

[4] E. W. Scripture: *Tests of School Children*, *Educ. Rev.*, v, 1893, 52.

[5] J. McK. Cattell: *Tests of Senses and Faculties*, *Educ. Rev.*, v, 1893, 257.

or implicitly, many of the distinctions urged in our preceding chapter.[1] He raises the question pointedly: "What constitutes the difference between a psychological and an anthropometrical experiment?" and finds very valid differentiation. The psychological experiment assumes, on the part of the observer, "practice in introspection, in attentional concentration, and in control of the particular apparatus employed. The anthropometric measurement requires at most only enough practice to carry out instructions. The value of the former is as much qualitative (for analysis and description) as quantitative; the value of the latter is solely quantitative." Speaking further of the line between experimental psychology and anthropometry, he says, "It is as unfair to rob anthropometry of her experiments and dub them psychological, as it is to psychology to conceal the difficulty of her own experimentation by substituting the simpler anthropometrical for it." Though this article was written with the simple aim of making suggestions as to the psychological exhibit at Chicago, it had in it a wider significance, in implicitly drawing the line of demarcation between the exact introspective experimentation with trained observers which lies within psychology, and approximating measurements of mental capacity, made upon miscellaneous individuals, which lie outside psychology. If the validity of the distinction suggested in Titchener's article had been recognized, mental tests would not have been welcomed, as they were, as a new and promising psychological method. The distinction was not recognized, however, and mental tests for a time had their welcome in psychology. We may follow their fortunes further.

Gilbert, during 1893-4, made tests of the physical and mental development of New Haven (Conn.) school children.[2] These tests, he says, were "to aid in the analysis of mental phenomena,"—a purely psychological aim, surely. In December, 1895, the American Psychological Association appointed a committee, composed of Cattell, Baldwin, Jastrow, Sanford and Witmer, "to consider the feasibility of co-operation among the various psychological laboratories in the collection of mental and physical statistics."[3] Cattell and Farrand, in 1896, published a report of physical and mental tests on students of Columbia University.[5] In this article, they state the purpose of

---

[1] E. B. Titchener; *Anthropometry and Experimental Psychology*, *Philos. Rev.*, ii, 1893, 187.

[2] J. A. Gilbert: *Studies from Yale Psych. Lab.*, (Nov. 1894) II, 40-100. His mental tests measure muscle-sense, color-perception, suggestion, motor ability, fatigue, reaction-time, discrimination-time and i me memory.

[3] *Psych. Rev.*, iii, 1896, 122.

[4] *Ibid.*, 618.

such tests to be "the study of the development and correlation of mental and physical traits. . . . . The tests have two chief ends, the one genetic, the other quantitative. We wish to study growth as dependent on environment and heredity, and the correlation of traits from the point of view of exact science." The authors seem to have accepted something of the distinction urged by Titchener three years before, and to have in mind the limited psychological bearing of mental tests, for they refer to them as "anthropometric work." The reasons urged for the performance of tests are largely practical rather than scientific: they say that the tests are of interest and value to the individual; may make the psychological laboratory of value to the community in school tests, medical practice, etc.; and give advanced students practice in research. On the other hand, they find certain psychological ends for mental tests, since they are to study (1) the interrelation of the traits defined and measured, and (2) the development of the individual and the race. The first of these aims falls in with individual psychology, the second with genetic; and so far the authors seem to hold to the psychological possibilities of mental tests.

The committee of the association presented its preliminary report in December, 1896.[1] A second report, followed by discussion, was made at the meeting a year later.[2] In the former report, the purpose of mental tests is stated briefly; "to reveal individual differences and development." On the latter occasion, Jastrow gave a paper on "Popular Tests of Mental Capacity," which was supplemented by statements from Baldwin and Cattell. Jastrow was at pains to differentiate "the careful and ingenious analysis on the part of well trained and scientifically self-observant experimentalists," to which he says the main equipment of a psychological laboratory may wisely be devoted; and investigations "to establish the normal capacity of simple and typical sensory, motor, and intellectual endowments," in the average individual or in selected groups, and the distribution, development and correlation of such powers,—in short, tests or "mental anthropometry." Though in this statement, Jastrow distinctly separates tests and psychology, it is doubtful whether the distinction has been clearly recognized in psychological thought at large. Jastrow himself in his latest statement says:[3] "Call it mental anthropometry if you will, but do not disregard the valuable contributions to other divisions of psychology and to the general conceptions of this science which such investigation has the possibility, and in my opinion the probability, of contributing." He thus

[1] *Psych. Rev.*, iv, 1897, 132-138.
[2] *Ibid.*, v, 1898, 145, 146 and 172-179.
[3] *Ibid.*, viii, 1901, 14.

maintains the possibility of contributions to psychology from mental tests, as does Cattell in the last statement quoted from him. On the other hand, the fact that seven years have passed since the committee suggested its tests, and that psychologists have not made any considerable contributions in this field indicates that, whether consciously or not, the distinction between mental anthropometry and psychology has come to be accepted. Some tests have indeed been performed.[1] One of the investigators, Seashore, seems clear with regard to the distinction just referred to: "most of the development of a problem," he says, "can be carried on to the greatest advantage by the ordinary method with trained observers; it is mainly when we wish to determine the nature of the naïve experience or the uniformity of a certain tendency that we find it profitable to appeal to the statistical method."[2]

There are indications, indeed, that the practical purposes in mental tests are coming to be recognized as the real and only reason for their performance; and that the hope of making them contribute to psychology is being given up. The practical end has been recognized from the first by Galton, Cattell, Jastrow, and others, as a large reason for performing tests. It seems fair to say that the present attitude toward mental tests, in so far as there is such an attitude, emphasizes the practical and minimizes the psychological aspect. The fact that psychologists have gone into public schools with tests and urged practical benefits from their performance; the fact that, despite "possible and probable contributions" to psychology, one does not actually find that such contributions have been made;[3] the fact, finally, that the public schools themselves are instituting mental tests for practical purposes,[4] indicate that mental tests are finding their place as a form of anthropometry carried out on large bodies of individuals for practical ends, and not as

[1] J. A. Gilbert: *Researches on School Children and College Students*, *Univ. of Iowa Studies in Psych.*, i, 1897, 1; C. E. Seashore, *Some Psychological Statistics*, *ibid.*, ii, 1899, 1; E. A. Kirkpatrick, *Individual Tests of School Children*, *Psych. Rev.*, vii, 1900, 274.

[2] *Univ. of Iowa Studies in Psych.*, ii, 1899, 3.

[3] One of the latest contributions to mental tests is C. Wissler's *Correlation of Mental Tests and Physical Tests*, *Columbia University Dissertations*, *1901*. Wissler compiles the tests made for several years at Columbia by Cattell and others and concludes that (1) the laboratory mental tests show little inter-correlation; (2) the physical tests show a general tendency to correlate among themselves but only to a slight degree with the mental tests.

[4] The most striking example of this is the Department of Child Study and Pedagogic Investigation, of the Chicago Public Schools. *Cf.* its reports, 1898-99 ff. *Cf.* also A. MacDonald: *Report of U. S. Commissioner of Education*, 1897-8, 989 ff. and 1281 ff. Tests on a less extensive scale have been made in other schools.

a branch or method of pure psychology. If the discussion in this paper helps to make the distinction between psychology and mental anthropometry more explicit, and calls attention to the emphatically practical purposes of the latter, one of its aims will have been accomplished.

The present paper is an attempt to describe, criticise, and arrange tests of audition, which will admit of application by unskilled investigators. It recognizes the practical purposes of tests and formulates testing methods for extra-psychological ends only. It is not a contribution to psychological science, but to an application of psychology, mental anthropometry. This may itself in time become a scientific discipline; and this paper is offered as a modest contribution toward that end. As already stated, the auditory tests which will be considered are of two classes, anthropometrical and diagnostic. The former fall into two divisions: those of general auditory acuity, and those of musical capacity. These will be treated in Chapter III and Chapter IV respectively. Diagnostic tests occupy the concluding chapter, Chapter V.

CHAPTER III.

Tests of Acuity of Hearing.

This chapter falls into two parts: (a) tests which employ speech in examining acuity of hearing; and (b) tests of acuity with mechanically produced tones and noises.

A. Speech Tests of Acuity of Hearing.

The most important function of the human ear is the hearing of conversational speech. No examination of the auditory organ is entirely adequate unless it determines the practical efficiency of this capacity. Impaired acuity of hearing ordinarily first manifests itself by inability to hear speech distinctly and easily. Further, in some aural diseases, the ear no longer hears certain vocal elements, *e. g.*, the low, deep sounds of speech, or words containing them. Tests with spoken words may be employed to disclose either of these disturbances. Speech tests for diagnostic purposes, *i. e.*, to determine those vocal elements which have become imperceptible through disease and to interpret their loss as indicative of certain aural diseases, will be considered in a later section, under diagnostic tests. Here we are concerned only with tests employing speech to measure the degree of general auditory acuity. In such tests, words are commonly spoken at various distances from an individual and the maximal range is determined at which he hears the words correctly.

The reasons for employing speech tests may be briefly stated:

(1) They measure directly and unequivocally the most important function of the ear,—the hearing of conversational speech.

(2) Other tests do not measure this functional capacity. The results of tests with the watch, acoumeter, audiometer, tuning forks, and other mechanical sources of sound may not be considered an unequivocal measure of the individual's ability to hear speech.[1] Thus, one may hear the watch at a considerable range, and yet be relatively deaf for speech, or conversely. The ratio between the range for speech and that for a single tone, or noise, is not necessarily constant, even in the same individual. The reason for this is evident. Hearing a watch or acoumeter click, or a single tone, involves the perception of a limited number of auditory qualities, and, assuming the Helmholtz theory of audition to be correct, the normal functioning of a restricted and constant part of the basilar membrane. Perception of speech, on the other hand, requires the hearing of a large number of tonal qualities and noises of high, low and medium pitch, and, accordingly, unimpaired physiological functioning over a wide stretch of the basilar membrane. Further, defects in other parts of the auditory organ, as in the drum, or the chain of ossicles, might accommodate the transference of sound of one pitch, say that of the watch, and prevent the passage of certain sounds in the wider range of speech. For these reasons, speech tests alone are entirely reliable in determining the functioning condition of the ear for practical usefulness in life. Accordingly, while conditions will often dictate the use of the mechanical tests of hearing, explained in the next section, speech tests should, where possible, be given the preference.

The problem of arranging speech tests is as follows. The syllables, words, or sentences to be used in the tests must be chosen; the conditions under which the tests are to be applied must be stated; a large number of individuals must be tested and the normal range for speech perception determined, so that the individual's range, and hence his acuity of hearing, can be interpreted in terms of the norm. This problem is more complicated than it seems at first to be.

The selection of speech tests is difficult, first of all, because of the great variety and complexity of sounds involved in human speech. The elements of spoken speech vary in pitch, intensity, complexity, and clang-tint or 'quality.' Some elements, the vowels, are approximately pure clangs with faint

[1] *Diseases of the Ear*, D. B. St. J. Rosa, New York, 1885, 49-50. *Diagnosis and Treatment of Ear Diseases*, Buch, New York, 1880, 17. *Ohrenheilkunde*, A. Politzer, Stuttgart, 1893, 115, 119. *Schuluntersuchungen ü. d. kindliche Gehörorgan*, F. Bezold, Wiesbaden, 1885, 9.

traces of noise-elements; others, the consonants, are noises with more or less slight admixture of tonal elements.[1] As spoken in words and sentences, speech presents further complications due to accent, emphasis and inflection. The classical investigation of speech in its relation to audition is that of Oscar Wolf.[2] Its most important feature is a statement (*a*) of the pitches of the elements of speech and (*b*) of their intensities. (*a*) *The pitch of speech-elements.* Wolf accepted the pitch determinations for vowels made by Helmholtz, and made similar determinations for all consonants having vowel elements. The results are expressed in musical symbols or vibration rates, *e. g.*, the vowel "a" has a pitch of b^2 or 896 vibrations; the consonant "b," a pitch of e^1 or 320 vibrations. The various vocal elements extend in all over eight octaves, from the lowest sound, lingual-R which contains a tone of 16 vibrations, to the highest, the S-sound, with a tone of c^5 or 4,032 vibrations. In Wolf's classification, the consonants h, l, m, n and w are assigned no pitch, as they are in themselves devoid of characteristic tone, and assume that of the vowel to which they are associated. (*b*) *The intensity of speech-elements.* Wolf measured this by the comparative distances over which the various elements of speech carry and are still audible.[3] The vowel "a" (as in father) has the greatest range, 252 m.; the aspirate "h," the shortest range, 8.2 m. The other sounds fall between these limits, the vowels having much longer ranges than the consonants. The vowels, indeed, may be distinguished at distances at which the consonants are all inaudible. The consonants, further, vary very widely in their relative ranges. These determinations of Wolf, which furnish a complete inventory of the pitches and intensities of the elements of speech, have made possible its accurate application in auditory tests. Later investigation will doubtless set aside parts of Wolf's work, particularly the pitches assigned to the various vocal elements. L. Hermann,[4] for example, has analyzed the vowels phono-

[1] W. Wundt: *Grundzüge d. phys. Psych.*, ii, 1893, 49.

[2] *Sprache u. Ohr*, O. Wolf, Braunschweig, 1871; and various articles in otological journals, particularly, *Arch. f. Augen- u. Ohrenheilkunde*," III, Abth. 2, 35; and IV, Abth. 1, 125; and *Zeits. f. Ohrenheilkunde*, XX.

[3] C. Blake made a mechanical analysis of the relative intensity of consonants (*Zeitsch. f. Ohrenheilk.*, XI, 1 Heft). His results do not agree very well with those of Wolf. The latter's results are certainly to be followed in making speech tests, since he measured intensity by the relative distance at which the speech elements are heard by the human ear.

[4] *Pflüger's Archiv*, xlviii, 371 and liii, 1. Summary in Wundt, *op. cit.*, II, 49, 51, where a table compares Helmholtz' and Hermann's results. See also Helmholtz, *Sensations of Tone*, 1895, 109, for a comparative table of the pitches assigned the vowels by various investigators.

photometrically, and reached results which vary quite considerably from those of Helmholtz, whom Wolf follows in the pitches of vowels. The vowel clangs are found to be richer in distinguishable over-tones, than with Wolf, and variations in the absolute pitch of tonal elements are discovered. Recent work by Scripture[1] calls attention to the fact that the pitches are nearly always changing, and have values that vary with the interconnection of syllables, and the turns of expression given the words.[2] On the other hand, Bezold reaches results confirmatory of the earlier work of Helmholtz and Wolf as to the pitches of vowels and consonants. Doubtless, there will be restatements of these values, perhaps in terms of the variable pitches to which Scripture's results call attention, before final agreement is reached; yet for the purposes of tests one is safe in following Wolf and the later otologists who have developed and applied his results. In our present enquiry, the selection of speech tests, the following facts, established by Wolf, are important:

(1) Vowels carry farther than consonants.[3] Words with the vowels of light clang-tint,—a, e, i,—are more easily understood than those with the vowels of dark clang-tint,—o and u.[4] (German sounds).

(2) The consonants h, l, m, n, and w may be disregarded in speech tests, since they have no tones of themselves, but depend upon the vowel preceding or following.[5]

(3) The consonant sounds to which special regard must be paid in making tests are grouped by Wolf as follows. The range of each element in loud speech is also given:

I group:	High and strong:	Sch,	200 paces.
		S,	175 paces.
		Soft-G, or Ch-soft,	130 paces.
	High and weak:	F,	67 paces.
II group:	Medium pitch, the explosives.		
		K and T,	63 paces.
		B,	41 paces.
III group:	Deep sounds.		
		Lingual-R,	41 paces.
		[U (whispered, in diagnosis)	50 paces.]

[1] E. W. Scripture: *Studies of Yale Psych. Lab.*, 1899 and 1902; and *Philosoph. Studien*, xix, 599.

[2] *Zeits. f. Ohrenheilk.*, xxx, 1897, 114.

[3] Stated also by Helmholtz, *op. cit.*, 68.

[4] Wolf: *Sprache u. Ohr*, 96.

[5] Wolf: *op. cit.*, 15. Wolf's statement may be overstrong. In words, these letters of course take on pitch values, so that Scripture assigns them pitches as other sounds. Doubtless, too, they have some tonal elements of themselves.

These are the three characteristic groups of consonants in Wolf's system. Any adequate speech test must examine the perception for sounds in all three classes. Since the vowels are of farther range than the consonants, testing is based on the hearing of consonant sounds. Regard need be paid to the vowels only to include those of dark as well as of light clang-tint. There are other considerations in addition to the necessary inclusion of consonants and vowels of the classes just mentioned, which determine the selection of the actual words or other speech material to be used in the tests.

This material might include (*a*) the sounds of the letters pronounced separately, (*b*) these sounds joined into nonsense syllables, (*c*) sounds joined into words, or finally (*d*) words joined into sentences. Wolf found the simple unconnected sounds of the letters desirable in diagnostic tests: of these we shall speak later. In the testing of acuity, which we are considering here, sentences are objectionable, since the person tested may hear certain of the sounds or words, and then, aided by inflection and emphasis, only too easily guess at the whole expression. Words are open to the same objection, though in a very slight degree, since inflection, emphasis, and context are lacking. Nonsense syllables and the disconnected sounds of the letters avoid this difficulty, but they are open to the weightier objection that they do not directly test hearing for conversational speech,—our aim in this determination. All things considered, single disparate words recommend themselves as the material for the tests.

Granted that words should form our testing material, there is the further problem of the particular words to be chosen. In such a choice, regard must be paid to two conditions: (1) the structure of the words, *i. e.*, the vocal sounds included in the words and the interconnections in which the elementary sounds stand; and (2) the meaning of the words chosen. Under the first condition, provision must be made, as we have seen, for consonants of Wolf's three classes, and as far as possible all other elements of speech, including vowels of dark and light clang-tint. In addition, the following minor points[1] should be observed as far as possible: i. Words beginning with two consonants present difficulties in defective hearing and may well be included in the series of test words. ii. Perception for a consonant sound varies slightly, in some cases considerably, according as it appears as the initial, final, or emphatic syllable in a word. This point is of importance in diagnostic tests rather than in this general test; nevertheless, words with consonants in a variety of positions are desirable. iii. A regular

[1] Wolf: *Sprache u. Ohr.*, 95-96.

alternation of vowels and consonants in a word facilitates its perception.

The second condition limiting choice of test words, that of their meaning, is based on the fact that the ease of perception of a word varies directly with its familiarity. This is a very important factor. In ordinary conversation, we apperceive much more than we actually hear.[1] As already stated, this makes sentences undesirable for testing. With single words, even, it is unavoidable that to some indeterminate extent the test word will be apperceived when only part of its constituent sounds are heard. The determining factor in apperception of a disparate word from partial hearing is the degree of familiarity of the hearer with the word. An illustration will show where this fact leads us. The word "electricity," for example, other things equal, will be perceived accurately at a greater distance by a student of physics than the word "calico." The former word fits into his most common thoughts, and the least auditory suggestion of its sound is sufficient to evoke the verbal idea. The latter word, on the other hand, is an unusual one in his circle of thought, and a distinct perception of its auditory symbol is necessary to call the verbal idea to mind. To assume to compare the range of hearing for one word with that for another, as is necessary in testing a single individual, it is evident that we must have a series of test words with which the person tested is uniformly familiar, words whose sounds are equal in apperceptive value, *i. e.*, in calling up the verbal ideas corresponding to them. Let us follow our illustration farther. While a student of physics would understand the word "electricity" at a greater distance than the word "calico," a dry goods clerk would reverse the relative facility of apperception, and hear the word "calico" relatively farther than the word "electricity." As it is necessary to secure comparable results from various individuals, in order to determine the normal range for hearing and evaluate individual variations from it, test words must be uniformly familiar not only to the single individual, but, as nearly as possible, uniformly familiar to all individuals tested. This second factor, that of uniform apperceptive value, might be satisfied in either of two ways. (I) In a negative way, by taking vocal sounds entirely unfamiliar to the person tested; *e. g.*, the individual speech sounds given disparately, or nonsense syllables, or perhaps rare English or foreign words. The disadvantages of this are obvious: the tests would rapidly become unpleasant; unmeaning vocal combinations, whether in nonsense-syllables or unheard-of words,

[1] W. C. Bagley: *The Apperception of the Spoken Sentence, American Journal of Psychology*, xii, 80.

are intrinsically difficult to understand, for in ordinary life one pronounces a new word two or three times before feeling sure it is understood correctly; finally, such tests are too far removed from practical hearing to measure the latter adequately. (II) In a positive way, uniform apperceptive value may be sought by taking words in common use, known to every one, and, as far as may be, equally well known to every one. Number-words, the names of the numerals one to ninety-nine or higher, have been suggested as meeting this requirement.

Number-words present three distinct advantages which recommend them for testing purposes: (I) As a group of words, number-words are doubtless the most uniformly known, with any individual, of any series that could be devised. 52 is as familiar as 26; 13 as 8, etc. (II) Of all words, the number-words are doubtless the most uniformly familiar from person to person, despite differences of age, occupation, etc. (III) They are understood by children easily and with practical uniformity. They are the first body of words systematically taught, and are of daily use in school. Hence number-words are especially suitable for testing the hearing of school children, even in the lowest grades. Allowances must, of course, be made in actual tests for younger children who can count to ten or twenty only.

In view of these advantages, which ensure comparable results in the individuals, and between individuals as well, number-words commend themselves strongly as the material for tests. This statement, however, looks only to the second condition of choice, the apperceptive value of words. Do number-words satisfy as well the first criterion; do they include the various elementary sounds, both vowels and consonants, in sufficient number and variety of arrangement, to give adequate testing material? The series, nine to ninety-nine, is obviously limited. There are within it twenty-seven primary words which, either standing alone as "five" and "thirteen," or forming double words, as "twenty-seven" and "fifty-six," give the ninety-nine different number words in the series. It will be to the point to examine the primary number-words with regard to their consonantal and vowel sounds. Webster[1] gives the following pronunciation, to which in some cases are added German equivalents[2] for convenience of reference to Wolf's standards:

1.	wŭn		6.	sĭks	
2.	tōō	ōō = approx. Ger. ū	7.	sĕv'n	
3.	thrē	ē = Ger. i or ie	8.	āt	ā = Ger. ē
4.	fōr		9.	nīn	
5.	fīv	v = Ger. w; ī = Ger. ei, ai	10.	tĕn	

[1] *International Dictionary*, 1903.
[2] C. H. Grandgent: *German and English Sounds*, Boston, 1892.

11.	ėlĕv'n	20.	twĕntў ў = Ger. ĭ, also Eng. ĭ.
12.	twĕlv		
13.	thērtēn	30.	thērtў
14.	fōrtēn	40.	fôrtў
15.	fĭftēn	50.	fĭftў
16.	sĭkstēn	60.	sĭkstў
17.	sĕv'ntēn	70.	sĕv'ntў
18.	ātēn	80.	ātў
19.	nīntēn	90.	nīntў

An examination of these words shows the presence of the following speech elements, the accompanying numerals indicating the number of times each element appears:

Vowels

a	3 ā
e	8 ē, 7 ĕ, 1 ė, 2 ẽ
i	4 ī, 5 ĭ, 8 ў
o	2 ō, 1 ô
u	1 ŭ
oo	1 ōō

Consonants

8 f, 3 k, 2 l, 20 n, 6 r, 9 s, 20 t, 6 v, 3 w, 3 th (as in *thin*).

Thus there are included in the number series ten different consonantal speech elements, each present from two or three times to twenty times. Webster enumerates twenty-five consonant elements as belonging to the English speech. Those not included in the number series are: ch, b, d, g, h, j, l, m, ng, p, sh, th (as in *thy*), y, z, zh. The fact that the number series includes less than half the consonantal elements of speech seems at first to condemn it, as inadequate to form the basis of speech tests. The answer again is not so easy; the adequacy cannot be decided off-hand by a mere enumeration of the sounds present. If these sounds include types of all consonant elements, they may very well test adequately the hearing for all. Comparing the consonants of the number series with the classification of consonant elements given in Webster,[1] one notes that all classes of consonants find representation by one or more sounds in the number series except the momentary surds (b, d, j, g). The final decision, however, whether the number series has sufficient representatives of all types of consonants, must be based, not on a classification like that of Webster's, made with regard to the pronunciation of speech, but on a classification made with regard to the hearing of speech. The consonants of the number words must be referred to Wolf's

[1] *Op. cit.*, lxvi.

three-fold classification, and the adequacy of the number series judged by it. The following Table shows under which of Wolf's classes each number-word falls; if the word contains consonant sounds of more than one class, it is included under each such class:

LOW.	MEDIUM.	HIGH.
2	2	
3		
4		4
		5
	6	6
		7
	8	
	10	
	12	
13	13	13
14	14	14
	15	15
	16	16
	17	17
	18	
	19	
	20	
30	30	
40	40	40
	50	50
	60	60
	70	70
	80	
	90	

Three number-words, one, nine and eleven, do not contain consonants falling within Wolf's classification. This does not, of course, exclude them from use in the tests. Considering the words falling within the three classes, the primary number-words are found to be distributed among the three classes as follows:

	LOW.	MEDIUM.	HIGH.
No. of words exclusively in each class	1	6	2
No. of words in each class and in another class as well	3	8	8
No. of words in all three classes	3	3	3
No. of different words containing sounds of each class	7	20	13

Our examination of the twenty-seven primary words shows that they are well distributed in the three classes of consonants. With these primary words, are to be included, in selecting test

words, the double number-words made from the primaries, *e. g.*, thirty-one, fifty-six. The primary words and the double words give in all 96 words, exclusive of one, nine, eleven, which are at our disposal. This certainly will furnish adequate material for series of testing words, with few repetitions save as the primary number-words reappear, differently compounded, in the various double-words.

At this point, an objection that has been made against the use of number-words in tests must be considered. Politzer[1] urges that, as the tests involve the use of a few primary words, the numbers will be judged with increasing ease as the test progresses, and the subject will come to guess the words, by hearing the vowels only. This objection Politzer considers insuperable. It is answered, however, by experimental results. Bezold used number-words in a test of 1,981 school children and in other extensive tests, and employs them constantly in otological practice. On the basis of the results of his school test, he says that there is no evidence of increased ease in the perception of the words as the tests progress, and whatever error there is from this source is practically uniform, since all the subjects know that they are being tested with number-words.[2] The experience of the author indicates that practice is a considerable factor, at first, but one that quickly reaches its maximum. Politzer is doubtless right in his objection, but he overestimates its importance. The error is known in direction, fairly constant in different individuals, and from its quick rise and small magnitude practicably negligible. It is certainly preferable to the variable errors of unknown magnitude involved in the apperception of the miscellaneous words which we must use in case we discard number-words. It may be added that number-words are in common and successful use by European otologists.[3] The objection of Politzer may, therefore, be overruled.

We return to the problem of selecting test words. The next step is to arrange lists of number-words, each list to contain words representing all three groups of consonants, and as far as may be with vowels of both dark and light clang-tint, and with consonants in the various positions, initial, medium, and final. The following lists seem to meet these requirements. The lists, further, may be expected to give comparable results, since they are equivalent as regards the elementary number-words included in each, and as regards the relative prominence of the three classes of consonants in each.[4]

---

[1] Politzer: *op. cit.*, 117.

[2] Bezold: *Das kindliche Gehörorgan*, 5.

[3] Bezold: *Functionelle Prüfung*, 206.

[4] Each of the ten lists contains uniformly the following elementary number-syllables: 1, "one;" 1, "two;" 1, "three;" 2, "four's;"

I	II	III	IV	V	VI	VII	VIII	IX	X
6	84	19	90	25	14	8	52	73	24
29	69	53	7	13	31	93	35	41	95
42	17	34	39	46	9	27	64	16	62
87	92	28	62	7	65	60	81	95	49
53	33	97	84	54	98	15	6	57	80
94	26	45	21	70	76	74	19	38	71
70	50	72	56	91	40	36	78	20	16
35	75	60	75	83	23	49	40	89	3
18	48	3	43	68	52	82	23	64	58
61	1	86	18	92	87	51	97	2	37

Speech tests are ordinarily given in whisper tones. In cases of hard-hearing, and sometimes in ordinary testing, conversational tones or even louder tones are employed. The merits of these two degrees of voice intensity are as follows.

Conversational speech (1) measures unequivocally the auditory functioning of practical life. Infrequent cases occur in otological practice in which whisper speech is heard farther than conversation (Burkhardt-Merian); and the whisper range is not a definite function of the range for conversation (Politzer). (2) Conversational speech must be used in determining cases of hard-hearing.

Whisper-speech (1) has approximately one-third the range of conversation and so permits testing in more restricted quarters. (2) In whisper-speech the vowels are reduced in intensity, while the consonants are little changed, so that the intensity of the vocal elements is more uniform than in conversation. This reduces the likelihood that the words will be guessed when only the vowel sounds are heard. On this score, whisper-speech has a distinct advantage, particularly when a limited number of test-words, as the number-series, is used. (3) Whisper-speech practically tests hearing for conversation. For these reasons, whisper tests are desirable and generally adequate. Limits of space, indeed, will usually preclude the use of conversational speech tests.

2, "five's;" 2, "six's;" 2, "seven's;" 2, "eight's;" 2, "nine's;" 1, "thir—;" 1, "—teen;" 1, "twent—;" 8, "—ty's." The relative prominence of the three classes of consonants in the ten lists, expressed as ratios of low consonants to medium to high (*i. e.*, l: m: h), is as follows:

I, 4:11:8; II, 5:12:8; III, 5:11:8; IV, 5:11:8; V, 5:12:9; VI, 5:12:8; VII, 5:12:8; VIII, 5:12:8; IX, 6:10:8; X, 5:11:9.

The writer compared these frequencies of the number series with that obtaining in the miscellaneous words of conversational speech and literary English. These latter showed the three classes of consonants present in the ratio 6:12:8, which is practically identical with the ratios given by the lists of number-words above.

Before considering actual methods of testing, the following important preliminary points of procedure may be stated.

1. Words are more easily perceived when the hearer is accustomed to the speaker's voice. This desirable familiarity will be secured by the preliminary explanations and statements made regarding the tests, which should be supplemented by a trial series of whispered test-words given according to the method of the actual tests. This serves the further purpose of making the subject feel at ease during the testing. This last is a very important point with children, and scarcely less so with many older people.

2. Two seconds or thereabouts before each test word, a warning signal should be given to insure maximal attention when the word is pronounced. Some uniform mechanical noise, as the sound of a pencil struck on wood, makes a satisfactory signal. The customary word "ready," pronounced in loud tones, may be used; but this necessitates an adjustment of the vocal organs for conversational speech just before each whispered test word, and is likely to interfere with uniform intensity in the latter. The signal should be repeated to indicate that the word has been spoken.

3. In testing the hearing for one ear, this ear is turned toward the examiner and the other ear closed. This can be most conveniently accomplished by pressing the tragus with the finger against the mouth of the meatus. Or, stoppers may be employed: Wolf used a rubber stopper; Politzer, the moistened finger in the meatus, or an olive-shaped obturator; and Titchener recommends the cap-shaped rubber eraser (found on many lead pencils), filled with laboratory wax before using. Small corks softened with vaseline are also quite satisfactory. When loud words are spoken directly before a diseased ear in hard-hearing tests, one must be sure, when words are heard, that the ear on the other side is not functioning despite its closure. Dennert suggested the check test of stopping both ears and repeating the words before the diseased ear. If words are still heard, the results of the preceding tests are unreliable, as the perception in both tests was doubtless mediated by the other ear; if words are not heard with both ears closed, the former test of the diseased ear was reliable.

4. The subject must not see the lips of the examiner in the tests, as slightly deaf persons often unconsciously "read the lips." In tests of one ear, the subject sits sidewise to the examiner and his position in part guards against this error. It is safer to have the eyes closed or shielded; the mouth should be closed through the tests, as the intensity of sound varies according as the mouth is open or closed. Such precautions

should be observed constantly. Any intermission interferes with regular attention.

5. With children, special care is necessary. As suggested above, conversation about the test and friendly treatment will assist in putting them at ease. If inattention and restlessness intervene, the tests should be stopped and resumed later.

6. As far as possible, variations in the voice of the examiner must be avoided. An approximately even intensity may be secured by always pronouncing test words with the residue of air in the lungs after the normal expiration (Bezold). If the test words are given successively, there may be interpolated after each word a definite number of breathings, three or four, the following word being pronounced after the last expiration in the group of breathings. Regard must also be paid to securing uniform tempo and distinct articulation. It is impossible to rule out differences in clang-tint between different voices. The variations in the individual voice may be largely obviated by the precautions suggested, and with practice it attains a closer approximation to uniformity.

7. It is generally desirable to examine each ear separately and then to test binaural hearing. The binaural range (*a*) is often unrelated to those of the ears functioning separately; and (*b*) it is precisely this range which is of importance in daily life. In the binaural test, the subject faces the examiner and closes his eyes.

8. The testing room for whisper speech should have a range of 30 m., if possible. Tests may be made, however, with an extreme range of only 18-20 meters, and with the second method suggested below even shorter ranges are satisfactory. The normal whisper range given by various investigators, varies from 17-41 m.[1] These different values are doubtless ascribable to the use of whispers of different intensities, and to the varying acoustic conditions in the various tests. The acoustical error might be obviated by making tests in the open air; but such testing is not commonly practicable. By performing tests in rooms of average oblong shape, and noting dimensions and other matters of acoustical importance, as disturbing noises, investigators will add to the comparability of their results. Bezold found that bright sunlight streaming from a window across his testing room changed its acoustic properties. The room should be kept free from noise within and without. Uniformity

[1] The following ranges have been given as norms by various investigators: Hartmann and Siebenmann, 25-26 m. in a quiet room and 20 m. in a room disturbed by the noise of the day; O. Wolf, 20 m.; Chimani, with soldiers, 21 m. (Politzer, *op. cit.*, 116.) F. Bezold, in his tests of school children, 17-20 m; Matte and Schultes, testing soldiers with a medium strong whisper, 35-40 m.; *Zeits f. Phys. u. Psych.*, xvi, 308.

should be observed as to position of large pieces of furniture, opened and closed windows and doors, and other conditions affecting sound reflection. Tests for conversational speech require a range three times that for whisper-speech (Bezold), and so a minimum space of perhaps 50 m. A range-line should be chalked on the floor in the direction of the greater dimension of the room, passing through the center and scaled to meter and half-meter divisions. It may be supplemented, if necessary, by a meter stick for short ranges in hard-hearing. The examiner, in working up and down the range, should always face the subject squarely as he pronounces test words. The subject will have a constant position at the end of the range-line, where he should be seated comfortably. If the testing is by groups, the individuals may be seated along an arc of 15 m. radius cutting the range at right angles.

Actual Performance of Tests. If a group of subjects are to be examined, rough preliminary tests may be made by pronouncing number-words at various ranges, on the results of which the group may be divided into two or more sub-groups, each containing persons of about the same grade of acuity of hearing. This will facilitate the actual testing. The precautions already suggested are to be observed: familiarity with the speaker's voice and with whisper-speech, closure of the ear not under test, closing of subject's eyes and mouth, attention secured by the warning signal just before the test-word, uniformity in the examiner's voice, and constant acoustical conditions in the testing room. With the individual or group seated at the end of the range-line, the examiner pronounces successively in whisper-speech a series of ten number-words, at each of the ranges selected, following this procedure: The warning signal, or "rap," is given, followed at the conclusion of the next breath by the first test word, which is pronounced with the residual air in the lungs. It is followed by a second warning "rap," signifying to the subject that the word has been pronounced and that he may now open his eyes and write down the numerical symbol of the word heard, *e. g.*, 22 or 9. If he has heard nothing, he indicates this fact by a dash (——). Meantime, the examiner interpolates three complete breaths and then gives the warning signal preparatory to the second test-word. At the signal, the subject closes his eyes and sits at attention until the examiner at the end of the next breath gives the second test-word, indicating its conclusion by another signal. Thus there follow successively: three complete breathings, the preparatory signal, the test-word at the conclusion of the next breath, the concluding signal, and so on through the series of ten words. This is the general procedure in the tests. A further statement must be made of two methods, either of

which may be followed in the tests: (1) the common method of measuring auditory acuity by the extreme range of accurate hearing; (2) the author's method of measuring acuity in terms of degree of accuracy of hearing at some constant range or ranges.

Method of Extreme Ranges. In this method the examiner makes rough preliminary tests with a few whispered number-words to determine the subject's approximate extreme range. An accurate determination of the audible range is then made by beginning at a point on the range-line 3-4 meters beyond this approximate range and moving in by meter or half-meter intervals, giving a series of ten test-words at each point. The point at which the subject hears correctly some arbitrary part, say 80%, of the list of ten words, is regarded as the individual's extreme range. The average of the ranges of all the individuals tested under the same conditions is the normal range. The individual's auditory acuity is expressed as a fraction whose numerator is his own range and whose denominator is the average or normal range.

The propriety of this method depends on the assumption that the intensity of the sounds of speech decreases with approximate regularity as the distance from the speaker increases. The author's experimental use of speech tests has led him to doubt the truth of this assumption. As set forth in the appendix, his results indicate that in a closed room there is no regular scale of decreasing intensity as the range increases. There is a general decrease in intensity as the indoor range increases, but the reflection of sound-waves back and forth in the room makes almost meaningless such distinctions as that A hears 80% correctly at thirty feet, B at forty feet, and C at fifty feet. It may happen, as it did in the author's work, that it is paradoxically easier to hear words in the particular room at fifty feet than at forty feet. The existence of this source of error throws no little doubt, in the author's opinion, upon the work which has heretofore been done in speech tests. Further experimental testing is needed before final conclusions can be stated regarding the extent to which the above method is trustworthy. Meantime, the following second method of procedure is advanced with the design of avoiding the error due to reflection of sound-waves, which seems to be inherent in the above method.

Method of Degree of Accuracy. In this method, one, two, or more, constant ranges are selected according to the testing space at the disposal of the examiner. If a single range is used, it should be on the border line of difficult hearing; if three ranges are chosen, one may be at easy hearing distance, one at a moderately difficult distance, and one at a quite difficult distance.

All ten testing lists, *i. e.*, the whole hundred testing words, are then spoken at the single range; or if more than one range is used, the words are divided about equally between the two or three ranges. Each individual is examined with the 100 words, and his results are computed as to the percentage of accurate audition. The norm is taken as the average of all the individual percentages; and each individual's audition is then expressed as a fraction of this norm. If time is valuable, the test may be shortened to fifty words or even less. If space is restricted, the range may be shortened by the use of screens to cut off the sound-waves and so reduce the intensity of the sounds. This may be easily accomplished by the examiner taking his position in an adjoining room with the connecting door open. It is necessary with this method, and especially so with artificially shortened ranges, that the examiner and the subject keep certain constant positions, even with regard to the direction of the head, throughout the entire series of tests.

The advantage of this second method lies in its freedom from the errors due to the reflection of sound-waves. With the examiner and subject in certain definite positions at a certain constant range, or ranges, the acoustical conditions remain uniform throughout the tests; and if A hears 80% accurately and B 60% accurately, it may be concluded with apparent certainty that A's hearing is more acute than B's. If the average accuracy of all persons tested, moreover, is 70%, A may be regarded as of super-normal acuity, and B as of sub-normal acuity. The theoretical superiority of this method is supplemented, too, by certain experimental results which, while not furnishing final proof, do in certain instances point toward the correctness of this method. These are given in the appendix.

Summary of Speech Tests. This discussion of speech tests has urged their importance as the only unequivocal measure of the ability to hear speech. An analysis of speech into its elements (following Wolf), and an examination of the apperceptive conditions of audition, furnished the criteria by which the material for speech-tests was selected. Judged by these criteria, number-words were found to recommend themselves for use in tests. Ten comparable and equivalent lists of ten number-words each were selected. The general conditions of conducting speech tests were given, followed by a statement of two methods of testing: the first, the common method of measuring the individual's acuity in terms of his extreme range of audition; the second, a proposed method of pronouncing all the test-words at one or more common constant ranges, and determining the individual's acuity in terms of the percentage of accuracy of audition.

B. *Tests of Acuity of Hearing with Tones and Noises.*

While speech tests alone measure unequivocally the ability to hear speech and, therefore, are to be employed, if possible, to determine acuity of hearing, limitations of space and other difficulties of using speech tests will often require the employment of tests with mechanically produced sounds. Reserving for a later section those methods which have their particular significance in the diagnosis of aural dieases, we will consider here, I, the four chief methods of testing acuity of hearing with mechanical sounds,—(*a*) with the watch, (*b*) acoumeter, (*c*) audiometer and other electrical devices, and (*d*) with forks; II, the tests of the lower and upper limits of tonal audibility.

I. Tests of Acuity of Hearing. It is a familiar fact that individuals vary in sharpness of hearing. Hard-hearing is an almost invariable accompaniment of old age,[1] and many individuals, young and old, have sub-normal keenness of hearing through innate conditions, or as a result of disease. The disturbance ordinarily manifests itself in difficulty in conversation, though, if the deficiency is slight, the person affected and his friends may alike be ignorant of its existence. The detection of such cases is particularly important in the public schools, where pupils who are regarded as stupid are often only hard of hearing, and could be properly provided for by advantageous seating and other special attention, were the real nature of their trouble understood. Tests of auditory acuity should also be applied to railroad men, telegraph and telephone operators, the police, soldiers, and all individuals in public or private service of a nature that demands accurate audition. In part, speech tests can be used for the purpose; but tests more convenient of application are often demanded, even if their results are not so unequivocal. Further, tests of finer adjustment than speech tests are needed to detect slight reductions in acuity of hearing or slight improvement in cases under treatment.[2] The tests with mechanical tones and noises meet these requirements. We shall consider several such tests in order.

A. Watch Test. This has long been the most common test of hearing and is still the one most frequently used in medical practice. The objections to the watch test have already been partly set forth in treating speech tests. Its results are not a

---

[1] Bezold's results give a comparison of hearing in school-children, and in persons over fifty years of age, as follows: 46.5% of the school-children examined had a range for whisper-speech of 16 m. or more; only 15.57% of persons fifty to sixty years old had a range of 8 m. or more, and only 6.7% of those from sixty to seventy had this range. Not only is there a decrease in acuity in old age, but the decrease becomes relatively more rapid as it proceeds. *Funkt. Prüf.*, 152-154.

[2] Politzer: *op. cit.*, 110.

reliable indication of the capacity to perceive speech; its sounds are a series of equidistant tones (Wolf) and so give rise to a perception of rhythm, not of simple sound; its ticking is so familiar that illusions of hearing are very frequent, *i. e.*, the subject imagines he hears the watch and reports it as heard even when it is beyond range; watches vary in the quality and intensity of their sounds, and hence results are not directly comparable.[1] On the other hand, the watch is always at hand and convenient to employ, and the procedure can be so arranged as to resolve, in part, the objections just stated. Nevertheless, the objections stand in part, and the writer recommends strongly that the watch be used only in cases where more accurate and standardized instruments cannot be secured.[2]

The procedure for watch tests consists essentially in determining the last point at which the ticking can be heard as the watch is carried out from the ear; then, after passing beyond the audible range, the watch is brought in toward the ear until again heard; the average of the two distances at which it is just heard is the range for the ear tested. The details of procedure may be stated as follows:

1. A range should be chalked on the floor 9-10 meters long, marked with half-meter sub-divisions from the one-meter point on. The meter stick itself should be used for distances less than a meter, being held for this purpose at right angles to the ear, and the watch moved along it. The range should extend along the middle of the room, and, as in the speech tests, the subject should be seated.
2. The subject's eyes should be kept closed, or better, covered with a blind. If he carries a watch, it must be laid aside. Quiet must be observed within and about the room.
3. The watch should be carried out along a line at the level of the ear and a little forward of a line passing through the two ears. It should always be held in the same relative position, preferably with its face toward the ear tested.
4. Tests should be made at the half-meter points along the range till a point is reached at which the ticking is no longer heard. The procedure is then reversed, and the watch brought in by stages, until again heard. The average of the two values is the range for the ear tested.

---

[1] F. Bezold states fully the objections to the watch in *Archives of Otology*, xv, 1886, 79; and on p. 9 of *Das kindl. Gehörorgan* gives comparative results showing the discrepancies between watch, acoumeter and speech tests. The acoumeter and voice agree better than the watch and voice. *Cf.* Tafel I, opp. p. 94, *op. cit.*

[2] Cattell and Farrand used the stop-watch (*Psych. Rev.*, iii, 635). It is also recommended by the Committee on Tests (*ibid.*, iv, 134). An ordinary watch, however, will serve as well if a stop-watch is not at hand. Stop-watches have been suggested by some with the method of stopping and starting the watch and requiring the subject to tell how many times he hears the watch going. The click made in starting and stopping most stop-watches is so loud, however, that the method is useless. This method can be used with an ordinary watch by covering and uncovering it in the hands.

5. Check tests should be constantly interpolated, especially with children, by covering the watch with the two hands and then asking if it is heard. This will disclose illusory judgments.

6. Each ear is to be tested separately, the other being meantime securely closed.

7. A normal range must be determined for the particular watch used. This can be taken as the average of a large number of individual determinations, and will vary with different watches. The hearing acuity of an auditory organ is then expressed as a fraction, whose numerator is its range, and whose denominator is the normal range; *e. g.*, if the average range is 2 m. and A's left ear has a range of 1.5 m., its acuity is ¾ normal. Expressed in such a way, the results of different investigators using different watches can be compared.

B. Acoumeter Tests. Dissatisfaction with the watch as a means of testing led Politzer to invent a small instrument, the acoumeter, which could be used as a substitute.[1] It consists of a small steel cylinder fastened at right angles into a hard-rubber upright, with a percussion hammer which can be raised a constant distance and allowed to fall upon the cylinder. This gives a tone (c^2), uniform in pitch and intensity. The instrument as described is arranged for air-conduction. It is also provided with a metal disc which can be set directly against the temporal bone, mastoid process, or other parts of the skull, if it is desired to test hearing with bone-conduction. The general procedure with the acoumeter, and the conditions to be observed, are precisely the same as in the watch test, and in making tests, reference should be had to the statements given in the preceding sections. In addition, the following points should be observed.

A large room will be needed for testing, for the instrument has a normal range of 15 m., which extends for very keen ears to 20 m. or more. The acoumeter must always be held perpendicular in giving the stimulus. The "ready" signal is given before each test, and fatigue avoided by allowing a short interval between successive tests, and limiting the length of a series. The actual procedure, in the receding series, consists of giving a varying number of acoumeter tones (1, 2, 3 or 4) at the successive points on the range, 10 m., 10.5 m., etc. The subject each time reports the number of tones he hears. The receding range is that at which he last reports correctly the number of tones given. The approximating series is made in a similar way, by starting at a point at which the sounds are inaudible. The range at which there is the first accurate report of the number of tones given, is the approximating range. The two are averaged for the range of the ear under test. The normal range is taken as the average of the various individual ranges, as with the watch, and the acuity of an individual ear is expressed as a fraction whose numerator is the range for the particular ear, and whose denominator is the average range; thus the results of various investigators are made comparable.[2]

---

[1] Politzer: *op. cit.*, 107-8.

[2] Politzer and others give the normal acoumeter range as 15 m. Bezold's results (*Das kindl. Gehörorgan*, 7) indicate a normal range of 16 or 17 m. This emphasizes the necessity of computing the normal

The acoumeter gives a more reliable measure of relative acuity for hearing speech, than does the watch.[1] It is, therefore, to be recommended instead of the watch. Its range is approximately the same as that for whisper-speech, and accordingly it has no advantage over speech tests in the space required for its employment. It should be borne in mind that the ranges of hearing with the watch, acoumeter, audiometer, etc., have no psychological significance, and are not to be confused with psychophysical limens.

C. The Audiometer and other Electrical Devices. Many attempts have been made to adapt electrical devices as testing instruments. The general plan has been to use a telephone click which will vary in intensity according to the amount of electric energy employed in producing it. The amount of current so used has been varied by a sliding induction coil, or by adjustable resistance, and admits, of course, of ready measurement. This, in turn, evaluates the intensity of the sound produced. Hughes and Boudet, Hartmann, Blyth, Koerting, Urbantschitsch, Preyer and others have devised and used various instruments of this sort.[2] Politzer makes the general objection that such instruments do not provide accurate measurement of sound intensity. This difficulty is apparently resolved in an audiometer recently devised by Seashore.[3]

This instrument gives a double telephone click, whose intensity can be varied through forty stages which are approximately equal in consciousness. The instrument includes a dry battery, primary coil, variable resistance coil, variable secondary coil, and a galvanometer, with necessary switches and wires leading to a telephone receiver. The primary current, by means of the variable resistance, can be adjusted to remain at a certain constant amount. The induced current may be varied in intensity through the forty steps, which furnish a measure of the intensity of the click given in the receiver on making and breaking this current. This instrument can be used with limited floor space, and outside noises give less disturbance than with most testing methods. An ascending and a descending series of tests are made, and the results averaged. Check tests should be inserted, especially when near the limen of hearing, by giving the "ready" signal and then omitting the stimulus. The two sounds produced by the Seashore audiometer are not, at least not always, alike in quality. This objectionable feature could doubtless be eliminated by a modification of the apparatus. The results are in terms of the arbitrary scale of forty steps in intensity; but they can be transposed into a more comparable

---

(average) range for the particular instrument used, and the particular conditions of each test. Politzer calls attention (109) to variations in individual results on different days, at different hours of the day, and under different physical and mental conditions (fatigue, anxiety, inattention, etc.). These facts should be remembered in employing the acoumeter, and, indeed, in all testing.

[1] Politzer: *op. cit.*, 109; Bezold: *op. cit.*, Tafel I, opp. p. 94.

[2] Gruber: *Diseases of the Ear*, 131; Politzer: *op. cit.*, 110.

[3] For a full description, see *Univ. of Iowa Studies in Psych.*, II, 158. This instrument is sold by the Chicago Laboratory Supply Co.

and intelligible form by writing them as fractions of the normal (average) result given by all the individuals tested.[1] The audiometer is also arranged to give a tonal stimulus. The tone may be varied and measured, Seashore says, in the same way as the noise stimulus. The author has not had an opportunity to test the audiometer in practical work.

D. Acuity Tests with Forks. The tuning fork is not adapted for acuity tests based on the distance at which a tone can be heard, since the tone given by a fork sounds first at a high intensity and then gradually "dies off;" when first struck, accordingly, it can be heard at a greater distance than a moment later when its tone has partially rung off. Van Conta[2] proposed a temporal method for forks which compares the ringing-off time for the ear under test with that for another ear used as a standard. It is based on the fact that the greater the acuity of hearing, the longer the tone will be heard.

A tuning fork (c^2) is struck and brought before the ear to be tested and when no longer heard, is transferred to the examiner's ear and the time it continues to sound noted. If necessary the procedure is reversed: the fork is placed first before the examiner's ear, and then before the ear under test. The obvious error in the variable time consumed in transferring the fork is avoided by Urbantschitsch, who uses a T-shaped tube with one branch in the subject's and one in the examiner's ear. The fork is held before the free end of the tube. By the use of two stop-watches each person can indicate accurately the length of time the tone is heard. There remains the error due to varying initial intensities of the tone, which can be reduced, but not eliminated, by care in striking the fork. This test obviously has a limited application.

II. Tests of Upper and Lower Limits of Hearing. Tonal sensations are evoked by the impact of periodic sound-waves upon the organ of hearing. Air waves of less than a certain minimal frequency do not produce a tonal sensation; similarly, air waves of more than a certain maximal frequency do not give rise to tones. We shall consider tests measuring the individual's upper and lower limits of tonal sensation. Such tests cannot pretend to the nicety of method or accuracy of results aimed at in psychological determinations. It is possible, however, to make rapid approximate measurements which will reveal individual variation in these capacities.

A. The Lower Limit of Tonal Sensation. Two distinct paths have been followed in determinations of the lower limen: it has been approached by trained investigators with skilled observers, and all the niceties of scientific method; so we have

---

[1] *Child Study Report, No. 2, 1899-1900, Chicago Public Schools,* gives results from testing 5,706 pupils with the audiometer. *Cf.* pp. 60 ff. See also *Some Results of Hearing-Tests of Chicago School Children,* D. P. MacMillan, *Medicine,* April, 1902.

[2] *Archiv f. Ohrenheilk.,* I, 107.

the work of Helmholtz, Wundt, Preyer, Appunn, and others who have secured results varying from 8 to 28 vibrations as the lowest audible tone.[1] Measurement of the lower limen has also been used in otological practice and in tests by Bezold, Zwaardemaker, and others, upon patients and other persons who are without training in scientific observation, and who present all extremes of general intelligence. Our proposed test plainly falls in with the latter rather than the former class. Of the instruments used in lower liminal determinations, the lamella and the wire forks are the most practicable for widespread use in tests. Testing methods will be stated for these two instruments.

(*a*) The lamella consists of a long, thin, steel blade, 420 mm. by 12 mm. by 1 mm., with a metal disc at its upper end. It is clamped in a wooden vise which in turn is fastened to a table. The lamella has a scale marked along one face showing the point to which it should be inserted in the vise to give 4, 6, 8, 10, ... 24 vibrations per second. When clamped in position at the desired point on its scale, it is put into vibration by springing it from a position of equilibrium with the finger. Thus sound waves of from 4 to 24 vibrations a second can be produced. Overtones are eliminated by adjusting a band of cloth one-third of the distance from the upper end of the blade.

There are two sources of error. (I) An error of observation. The subject is likely to mistake the discrete puffs of air at the lower rates for tones (Bezold). Practice should be given with a few preliminary tones at the higher rates which are usually clearly distinguished as tones, and then with the lowest rate which will give discrete puffs of air; and the subject's attention called to their qualitative difference. (II) An objective error. The length of the lamella increases as the tone deepens, and hence the amplitude of vibration and the intensity of the tone increases (Zwaardemaker). This cannot have much weight, as the difficulty of perception increases as the tone deepens, and the facility due to increased intensity may be regarded as an offset to the greater difficulty of perception. The procedure with the lamella involves the usual ascending and descending determinations and an averaging of results.

(*b*) The Appunn wire forks are made of wire bent in the usual fork form, and provided with discs on the ends of the prongs. The series includes forks differing by two vibrations from eight to twenty-four vibrations, and by eight vibrations from twenty-four to fifty-six vibrations. These forks present the following advantages: approximate freedom from overtones;

---

[1] Helmholtz used organ pipes and a vibrating monochord: *Sensations of Tone*, 1895, 175-6. Preyer used reeds and a heavily weighted fork: *Die Grenzen d. Tonwahrnehmung*, Jena, 1876. Wundt used the heavy fork, the Appunn lamella or vibrating blade, and difference tones: *Grundzüge d. phys. Psych.*, 4th ed. I, 450; other investigators have used Appunn's wire forks.

tones loud enough to be audible; the initial amplitude is constant for each fork; they are easily handled, and present no serious difficulties for untrained subjects. The series should be applied in both orders, and an average struck between the results.

With the lamella, Zwaardemaker found average values for the lower limen for different ages at from 10 vibrations to 13 vibrations.[1] Bezold, working with loaded forks (Edelmann's), uses a lower limen of 16 vibrations, though for many persons, he says, this doubtless lies above the actual limen. The general results of tests will probably agree more closely with Zwaardemaker's figures.

B. The Upper Limit of Tonal Sensation. The most practical instrument for upper limit tests is the Galton graduated whistle. This was invented by Francis Galton and has been recently improved by Edelmann.[2] It consists of a covered cylindrical organ-pipe of small dimensions with a small air-bulb as bellows. The pitch of the whistle is varied by a micrometer screw which moves a piston up and down the cylinder, thus altering the length of the pipe. Readings on the micrometer scale correspond to the pitches to which the whistle can be adjusted. As modified by Edelmann, it has a second micrometer screw which adjusts the lips of the whistle to give an optimal tone at the various pitches. A schedule accompanies each instrument showing the proper setting of the two adjustments for the pitches within its compass. One form covers the tonal scale from a^4 of 3,480 vibrations to f^8 of 44,193 vibrations and above that to 49,000 vibrations. A scale of

---

[1] A. Zwaardemaker: *Zeitsch f. Psych. u. Phys.*, vii, 1894, 10. Zwaardemaker in this article attempts to establish a correlation between increasing age and a shortening of the tonal scale. The average lower limit for ten to twenty years is 10.10 vibrations; for thirty to forty years, 10.85 vibrations; over sixty years, 12.95 vibrations. Tonal audition extends over eleven octaves in youth, he concludes, and only ten octaves in old age. He reached these results with the lamella and the Galton whistle. Bezold (*Funkt. Prüfung*, 144), using the weighted forks and the whistle, finds that the shrinkage is much less, and explains it as due to injuries and disease, rather than advancing age.

[2] Determinations of the upper limit have also been made in the past with small forks, made by Appunn, Koenig, and Edelmann, and with Koenig's cylinders. Appunn's forks consist of one series of eleven forks, varying by steps of 5,000 vibrations, from 5,000 to 45,000, and another series of thirty-three forks from c^4 (2,408 vib.) to g^8 (49,152 vib.). Discredit has been thrown on determinations with Appunn's forks through recent investigations showing large errors in their vibration markings. (For statement of discussion see *Zeitsch. f. Psych. u. Phys.*, xxi, 141; xxii, 229; xxiv, 171 and 367.) Koenig's cylinders are of steel, 20 mm. in diameter, and of varying length. One set of ten cylinders covers the range from 4,096 to 32,768 vibrations. There is also another series of 22 cylinders covering the same interval.

intervals by steps of 1,000 vibrations is also usually given, from 10,000 vibrations to the upper limit. Zwaardemaker notes two advantages of the whistle: it is approximately constant in intensity,[1] and its tone is prolonged but an instant, and so does not fatigue the ear. Further, it may be said that the whistle is simple to operate, and gives a tone which when present is not difficult to distinguish from the accompanying noise. In testing with the instrument, an ascending and descending determination should be made and averaged. Checks should be made on judgments at the limen, by interpolating occasional pure noises between the tones. Results may be expected to vary considerably; it should be remembered that large differences in vibration rates at 30,000 to 50,000 vibrations mean only a small tonal difference; *e. g.*, the interval just stated, which comprises 20,000 vibrations, is only a major 6th.[2] Perception at the upper limen is very fatiguing, and series must accordingly be limited.

APPENDIX.

AN EXPERIMENTAL EXAMINATION OF SPEECH-METHODS.

We have given in the text what may be called the 'traditional' method of testing audition by means of speech. This method, it will be remembered (see p. 40), is the method of Extreme Range of Hearing. A normal maximal distance is selected and, on the basis of it, individual capacities are computed by comparing the individual ranges with the normal range. Thus, if the normal range is twenty meters and the individual's limit falls at ten meters, the individual's capacity is represented by some fraction: *e. g.*, ½.

Now it is evident that the correlation of auditory capacity with distance will be valid only on condition that distance and audibility stand to each other in some simple functional rela-

[1] E. W. Scripture: *Studies of Yale Psych. Lab.*, II, secured constant pressure by mechanical means. See, however, Myers' results in the next foot-note.

[2] Edelmann gives 50,000 vibrations, or g^8, as the upper limit; Schwendt, e^8-f^8 (41,000-44,000 vibs.); Zwaardemaker, with persons of all ages, from c^6 to f^7 (22,000 vibs.). For Edelmann, see *Zeitsch. f. Ohrenheilk.*, xxxvi, 1900, 330; Schwendt, *Pflüger's Archiv*, lxxv, 346, lxxvi, 189, and *Archiv f. Ohrenheilk.*, xlix, 1; Zwaardemaker, *Zeitsch. f. Psych. u. Phys.*, vii, 10. Very recent determinations by C. S. Myers (*Journal of Physiology*, London, xxviii, 1902, 417-425) throw doubt on all high results with the Galton whistle. Myers concludes that the pitch of the whistle varies with the absolute air pressure and with changes in air pressure; and that the liminal value of 50,000 vibs. is unreliable. He puts the limen at 20,000 to 25,000 vibrations for young adults. Tests may be expected to give results like Zwaardemaker's and Myers', *i. e.*, about 20,000 vibs.

tion, and further, on condition that the sounds used as test-words undergo, with change of distance, merely a quantitative, not a qualitative, alteration. These conditions have been assumed, without sufficient ground, by the traditional method.

The obvious way to test these assumptions is by appeal to experiment. Two groups of experiments were, accordingly, arranged and carried through in a manner presently to be described. In the first group, a large number of ranges was employed as a preliminary step to the selection of a final, maximal range, as provided for by the method. These experiments brought to light the fact that the two essential conditions just named are not to be found in the method as it stands.

The second group had, then, for its end the discovery of a more adequate mode of procedure.

I. *Preliminary Tests for the Determination of Extreme Range.* The experiments in Group I were conducted in the Armory Hall of Cornell University, a rectangular room 40 by 120 feet. The range began at a point twenty feet from one end of the room, and was laid off along the median line for a distance of eighty feet. The person giving the test-words stood at the zero point of the range, while the observers took up their positions, 20, 30, 40, . . . 80 feet, along the line indicated. A Table of tests, consisting of ten series of ten number-words each, was compiled.[1] These series were given at the various points along the line, and the subjects wrote down the numerical symbols of the words so far as they were heard. Results were computed on the basis of percentages of accuracy in each series, part credit being given where syllables, and not whole words, were heard correctly. There were ten observers. They are indicated in the Tables by the letters A, B, ... J.[2]

---

[1]These words (the original lists devised by the writer) were as follows:

I. 6, 42, 87, 34, 29, 61, 94, 18, 70, 53; II. 92, 33, 12, 75, 5, 27, 8, 69, 47, 50; III. 19, 28, 35, 60, 51, 86, 3, 98, 79, 45; IV. 96, 20, 41, 15, 89, 72, 57, 38, 2, 65; V. 13, 25, 46, 32, 7, 83, 68, 54, 73, 90; VI. 31, 4, 23, 14, 48, 67, 85, 91, 52, 76; VII. 37, 26, 49, 8, 66, 17, 74, 58, 93, 82; VIII. 81, 22, 13, 6, 30, 78, 64, 97, 44, 55; IX. 62, 43, 39, 84, 14, 56, 7, 99, 21, 77; X. 3, 40, 95, 36, 16, 59, 10, 63, 24, 71.

Like the revised lists given in the paper, these series are comparable as regards the relative frequency of low, medium, and high consonants (Wolf's classes). They are not exactly equivalent, however, in point of primary number-words present in each. After using them in the first part of the experiments, revised series (p.36), which met precisely this second criterion of equivalence, were substituted.

[2]The subjects, whose assistance the author takes this occasion of acknowledging, were Professor I. M. Bentley and Dr. J. W. Baird, of the Department of Psychology; Messrs. R. B. Waugh and W. A. Frayer, who had had some training as observers in psychological work; and Messrs. R. W. Palmer, L. E. Palmer, T. E. Faxon, G. L. Genung, C. L. Rand, and R. G. Marvin.

Table I gives the average (percentages) of accuracy for each individual at the various ranges at which he was tested. Each percentage is the average of accuracy for all the series used with a given individual and for a single distance. The figure in parenthesis after the percentages indicates the number of series on which the average is based.

TABLE I.

Range in feet.	*20*	*25*	*30*	*40*	*50*	*60*	*65*	*70*	*75*	*80*
Person.										
A.			75.0(3)	60.0(3)	51.6(3)		33.0(6)		32.5(2)	38.3(3)
B.	90.0(1)	95.0(1)	79.2(5)	73.7(4)	68.7(3)	61.1(1)	28.2(5)	55.4(3)		47.5(6)
C.	90.0(1)		66.2(4)	56.2(6)	66.2(1)	43.7(1)	35.0(6)	40.0(1)	38.7(2)	29.6(4)
D.			62.5(5)	56.1(5)	47.5(5)	52.5(2)	47.5(2)	42.5(2)		
E.	95.0(1)	72.5(1)	71.2(2)	47.5(3)	36.4(5)	46.2(1)	53.7(2)	42.5(1)		37.9(3)
F.	97.5(2)	55.0(1)	83.7(1)	65.0(1)	66.2(1)	66.2(1)		47.5(1)		41.2(1)
G.			45.2(5)	38.1(5)	46.6(6)	40.0(2)	33.7(2)			
H.	65.0(1)	55.0(2)	48.7(1)	52.5(1)	60.0(1)	37.5(1)		37.5(1)		42.5(1)
I.	83.7(3)		65.0(1)	60.0(1)	86.2(1)	27.5(1)		30.0(1)		32.5(1)
J. (deaf ear.)	23.7(1)		15.4(3)	2.5(1)	10.0(1)	5.0(1)		0.0(1)		0.0(1)
Totals. (A-I)	87.36	66.5	66.14	55.00	53.30	46.70	35.50	45.1	35.5	37.30
M. V.	10.2	21.2	12.7	9.4	11.9	10.0	10.1	7.6	4.3	8.2

This Table indicates, from the standpoint of individual results, that while there is a general decrease in audibility as the distance increases, yet the decrease is neither constant nor uniform. For example, the percentages for A and B show a fairly uniform decrease through the middle ranges; but there is an unexpected rise in the figures at the 80-foot range. This inequality is due, apparently, to reflection of sound-waves from the farther end of the room—a factor which makes results at the longer ranges of this particular room difficult of interpretation. The results indicate, further, that not only are the changes in intensity of sound not uniform as the range increases, but that over a considerable middle part of the range the intensity of the sounds may be approximately constant, so that differences in distance do not give corresponding differences in intensity. See, for example, the percentages of C and G at 30, 40 and 50 feet; or of E at 40-65 feet.

We are aware of the possibility of error in these conclusions due to the small number of cases considered; particularly, since the test-series used in this part of the experimental work lack perfect equivalence. Whatever influence might arise from either of these sources, however, would scarcely be felt in the

averages for observers A—I, which are given at the bottom of the Table.

The lack of uniformity, both in the individual and the general averages, suggests strongly the limitations—not to say the inadequacy—of the commonly accepted method of speech tests—the method of extreme range of accurate hearing. It shows that, under the given conditions, there is no simple relation between length of range and the goodness of hearing. Over and above these general results, the tests just considered throw some light upon the relative ease of perception of the various elementary number-words which make up our lists. Table II shows the percentage of accurate audition for each syllable at four of the ranges. The last column gives the percentage of right cases (sounds heard correctly) for each syllable at all distances. The total number of syllables recorded was twenty-five hundred.

TABLE II.

RANGE	30 Feet.		50 Feet.		65 Feet.		80 Feet.		All ranges.	
	WORD	% R.	WORD	% R.	WORD	% R.	WORD	% R.	WORD	% R.
1st	twen	96.4	six	87.9	six	86.0	six	92.3	six	90.4
2nd	six	94.9	-ty	80.0	seven	71.1	-ty	66.6	-ty	76.7
3rd	-ty	89.4	eight	77.5	-ty	64.9	two	61.9	seven	70.2
4th	eight	79.4	twen	71.4	two	57.1	seven	57.1	twen	68.0
5th	seven	79.2	seven	70.7	twen	54.5	twen	36.8	two	64.6
6th	nine	68.9	two	70.3	eight	41.3	three	25.0	eight	58.7
7th	two	66.6	-teen	48.1	nine	22.5	eight	22.8	-teen	38.6
8th	one	65.0	one	42.8	five	13.0	-teen	22.2	nine	37.8
9th	thir	58.8	three	40.9	three	12.5	four	19.0	one	34.2
10th	-teen	57.6	thir	38.8	-teen	11.7	five	16.29	thir	32.4
11th	five	47.8	five	38.6	four	6.9	thir	13.6	five	31.8
12th	four	45.6	nine	32.0	one	6.6	one	11.7	three	25.8
13th	three	25.9	four	19.6	thir	0.1	nine	9.6	four	22.8

In this Table, "one" is heard moderately well throughout the shorter ranges and poorly at the longer. "Two" holds a fairly constant position, at 60%-70%. "Four," which has a medium audibility at thirty feet, drops at the longer distances. "Five" holds its medium place up to fifty feet, but falls later. "Six," on the other hand, maintains a high place throughout. Similarly, "Seven" keeps a high position, though it drops off a little at the longest range.

One may say, in general, on the basis of the Table, that (1) there are distinct differences in the ease of audition for the different number-syllables; although (2) these differences are not constant throughout the various ranges. It is obvious, *e. g.*, that 'six' and 'ty' are better heard, at any one of the distances used, than 'four' and 'five;' but it is also obvious that the

advantage which one syllable has at, say thirty feet, may be transferred to another syllable at sixty-five or eighty feet.

This rise and fall of syllables in the Table is significant. If change in distance at which a sound is heard means simply alteration in intensity, then all words should suffer the same fate, as the range is increased; but if the phonetic elements undergo change by reason of re-enforcement and absorption, then we should expect to find precisely what, as a matter of fact, we do find in our results. We find that 'twen,' *e. g.*, falls from ninety-six to thirty-seven percentage of right cases where 'two' falls only from sixty-seven to sixty-two. The one syllable suffers rapid dissolution; the other is heard practically as well at eighty feet as at thirty. Other similar comparisons may be drawn.

We may say, in concluding this section, that the distribution of 'right' cases—whether we consider gross averages of series (Table I), or averages for individual phonetic elements (Table II)—shows that the physical conditions under which the experiments were made were too complex and too variable to give a definite and unambiguous measure of auditory acuity in terms of distance. Any test which, like ours and like speech tests in general, involves reflecting and absorbing surfaces complicates hopelessly both the absolute intensity and the qualitative integrity of the sounds employed and, therefore, violates conditions which we found at the beginning of this section to be indispensable to the success of the method.

II. *Method of Degree of Accuracy.* The root of the difficulty in the method just criticised lies in its attempting to measure audition in terms of a 'normal' distance where no 'normal' distance can be secured. It should, however, be possible to save a remnant of the traditional method by making—not distances[1]—but the number of right cases at *a given constant distance* the basis of computation. This modification we have attempted to carry through in the experiments now to be described.

The second group of experiments was performed, for the most part, in two suites of rooms in the psychological laboratory, at Cornell University. One suite, here designated "A," consisted of a lecture room 24 by 42 feet, adjoining which, by a door near one corner, was a small room 12 by 22 feet; the subjects were seated in the corner of the room diagonally opposite the door, and the range was laid off along the diagonal. The experimenter occupied three constant positions, two on the diagonal, giving a 20-foot and a 40-foot range, respectively,

[1] Various distances are used in the following experiments, but only by way of finding the most convenient conditions under which to work.

and a third position, ten feet away from the diagonal and within the smaller adjoining room. This gave three constant ranges of 20, 40, and 50 feet, in room A. The second suite, "B," was quite similar except in size; the large room was 42 by 18 feet, and the small room 8 by 20 feet. The range extended along the median line of the large room to a 20-foot and a 40-foot point, and from the latter at right angles through an open doorway to a 55-foot point in the smaller room. Room "B," had, accordingly, three constant ranges, 20, 40, and 55 feet. The two rooms, "A" and "B," from their size and the different location of the ranges, may be regarded as distinctly unlike as regards acoustical conditions.

The method consisted in submitting to the subjects a series of ten words at 20 feet, a second series at 40 feet, a third at 50 feet, and so on until all ten lists were given, distributed in each case as follows: three lists at 20 feet, three lists at 40 feet, and four lists at 50 or 55 feet (according as the tests were in room "A" or "B"). Only one ear at a time was used.

TABLE III.

			20 Feet.		40 Feet.		50 Feet in A. 55 Feet in B.	
Obs.	Room	Ear		M. V.		M. V.		M. V.
K	A	L	80.5	9.2	85.7	5.3	46.2	13.3
		R	66.1	2.7	83.3	2.2	43.3	5.0
	B	L	91.6	3.9	90.2	3.5	54.3	7.4
		R	91.1	6.3	83.3	5.0	66.6	8.3
			82.3	9.0	85.6	2.3	52.6	7.8
L	A	L	86.9	8.3	86.6	8.9	65.1	10.2
		R	86.1	2.6	90.0	6.7	56.6	9.1
	B	L	87.2	9.2	92.4	3.3	61.8	9.3
		R	94.4	5.2	95.5	2.9	73.3	8.3
			88.6	2.8	91.1	2.8	64.2	5.0
M	A	L	72.2	4.1	82.2	5.9	60.0	4.2
	B	L	90.0	4.4	85.0	5.5	57.9	5.4
			81.1	8.9	83.6	1.4	58.9	1.0

Table III includes results from two sets of experiments carried out in rooms "A" and "B." Reading the tables horizontally, one gets the percentage of correct cases for series of ten words each,[1] at a given range. The letters A and B in the

[1] Most of the results are based upon twenty series for each line; a few upon ten series.

Table designate the rooms used for the particular series of tests; K, L, and M, the observers.[1]

Assuming an acoustical difference in the two rooms, "A" and "B," we might expect, if our method is sufficiently delicate, to find some indication of it in comparing the results of similar tests given in the two rooms. Now the results do show higher average, range for range, for both the right and left ear, in room B than in room A. There are only two places in which B gives slightly lower averages than A, and there the range is 55 feet as against 50 feet in A. There is, accordingly, practical uniformity of higher averages for B than for A. This can only be accounted for as due to better acoustical conditions in room B. But it is fair to maintain that since the proposed method is precise enough to reflect in its results such acoustical differences, its delicacy, in actual tests of auditory capacity, is assured.[2]

But the most important issue of the method appears when we compare observer with observer within the same room and at the same range. Here we come to the real test of the method. The following Table gives the order, from worst to best, of the three observers and over the three ranges.

TABLE IV.

	Room	20	40	50 and 55
Left Ear	A	M, K, L	M, K, L	K, M, L
	B	L, M, K	M, K, L	K, M, L
Right Ear	A	K, L	K, L	K, L
	B	K, L	K, L	K, L

We find in the Table that, at the fifty and fifty-five foot ranges, the order for the observers is invariable. With both ears and in both rooms L hears with greatest and K with least acuteness. At the smaller distances, the order is not entirely constant. In considering the results, it must be borne in mind that no one of the observers is noticeably hard of hearing. We cannot expect, therefore, to discover large differences of capacity. It must be noted, further, that the small distances give percentages too near one hundred to be very significant. It is only where acuteness is considerably reduced—as it is in the last column—that we could expect to look for distinct and uniform differences. Where the method is used with marked im-

[1] These included the following Cornell students whose help is acknowledged with thanks: R. P. Butler, A. B. Truman, and O. Goehle.

[2] The fact that the forty-foot averages are higher than the twenty-foot is in striking confirmation of our contention regarding the effect of reflection upon intensity.

pairment of auditory capacity, we may expect still more striking and unambiguous results. To make sure of this, the writer tested three individuals, N, O and P,[1] whose respective ages were nineteen, thirty-five and sixty years and whose hearing was plainly different.

The experiments, which were carried out in two adjacent rooms in a dwelling, gave as an average of five series of ten words each and with a range of eighteen feet, the following percentages of right cases:

$N = 100 \pm 0$; $O = 80 \pm 7.3$; $P = 51 \pm 3$.

Even with so limited a number of results, we are safe in recommending the method of Degree of Accuracy. To be sure, it gives values which are valid only under the particular conditions of the experiments; but this is true, as well, of the method of maximal range which we have been obliged to discard as untrustworthy. Our own method, besides laying claim to vastly greater accuracy than the old method, is much more easily carried through, since it can be used in rooms of ordinary size,—the only necessity being sufficient space to give a fairly great number of incorrect answers. When once the proper range has been found, all other ranges should be discarded and the percentages reckoned on the basis of a single distance.[2]

[1] J. A., G. and C. T. Andrews.

[2] Experiments are under way in the Cornell Laboratory on a speech-method which will, if successful, eliminate entirely the factor of distance and provide a standard series of tests which can be employed without regard to local conditions.